Editorial

Preserving the spoken word: six issues in search of a digital solution

The author

Michael Seadle is Editor of *Library Hi Tech*. He is also Digital Services and Copyright Librarian at Michigan State University, East Lansing, USA. <seadle@mail.lib.msu.edu>

Keywords

Preservation, Sound, Conversion

Abstract

How can we digitize the spoken word in a way that maximizes its usefulness and accessibility over as long a period as possible? This editorial lists six issues: copyright and ownership; selection at both the collection and contents level; conversion, formats, and analog input; segmentation and structuring; metadata at both the bibliographic and contents level; and maintenance and refreshing.

Library Hi Tech
Volume 17 · Number 3 · 1999 · pp. 230–232

· ISSN 0737-8831

> But isn't that the cause of all the trouble? Words! We all have a world of things inside ourselves and each one of us has his own private world. How can we understand each other if the words I use have the sense and the value that I expect them to have, but whoever is listening to me inevitably thinks that those same words have a different sense and value, because of the private world he has inside himself too ... – Pirandello (1921), *Six Characters in Search of an Author*, Act I.

As Pirandello suggests, words cause trouble. The word "preservation" especially causes trouble when coupled with "digital". A significant number of intelligent and capable people jump up every time I make such a link and remind me that digital preservation is an oxymoron. I always agree, abandon the offending phrase, and try to pursue instead the detailed elements of common interest rather than the definitions that divide. For me the real question is not whether digitization preserves, but how we digitize the spoken word in a way that maximizes its usefulness and accessibility over as long a period as possible.

Sound in general has attracted more and more attention on the Web as streaming and compression algorithms make it practical for people with ordinary modems and phone-line connections to hear digital sound without interminable delays. Sound is in its own way as rich as images. It is a chronic problem of e-mail that the written word does not convey the same range of emotion and intent that we understand implicitly in speech. A significant number of interviews, oral histories, and speeches have been appearing on the Web, and libraries are beginning to get grants to add to them, often without actually having a clear set of ideas about how to do it because no "best practices" exist. This editorial describes some of the issues that I think should be part of a set of best practices:

- *Copyright.* This is the first issue because it limits what can be done with all others. US copyright law generally treats spoken words like text. Once the words have been "fixed" on a reasonably permanent medium such as a tape or computer disk, they have legal protection. Normally the speaker owns the copyright, though a corporation might be the owner in a work-for-hire situation, and the words of a US federal government employee speaking on federal government

Library Hi Tech
Volume 17 · Number 3 · 1999 · 230–232

business might go directly into the public domain.

For anything which is protected, a library must have the copyright owner's permission to create a digital derivative (who is not necessarily the owner of the physical source), and may have to pay a royalty. There can also be secondary rights which broadcasters own, and performance permissions for certain types of works. The only clearly safe materials are those created before 1923 and those with explicit permission.

- *Selection.* This issue has two levels. At the highest level, the decisions resemble those for all other kinds of library materials: does the work fit in the collection policy, does someone want it, is it affordable? But at a more detailed level selection becomes less like book-buying and more like archival selection. This is because many of the materials that libraries want to digitize do not come packaged and pre-edited. More often they are a jumble of questions, answers, conversation (in the case of oral history), or long sequences of introductions, speakers, questions-and-answers, and background sounds (as in speeches).

  The selection process involves not only deciding whether to get an item, but what the item consists of. For example, in an oral history is it a single question-and-answer sequence, a single session out of several, all the interviews with one subject, all the interviews by a single interviewer, all the interviews that relate to a particular oral history project? The latter may be too large to be a single item, and the single question-and-answer sequence may be too small – but of course this will vary with circumstances. There can also be parts of a recording which are legal to copy, such as a President's State of the Union speech (arguably a government document), while the commentary would not be legal to take without permission.
- *Conversion.* This involves deciding on a digital format, such as whether to use .wav files or a "raw" pulse code modulation (PCM) bitstream. The .wav files have headers which can be used for basic bibliographic and conversion information that remains permanently attached to the sound. This is a feature which librarians like and audio engineers appear to hate. Conversion also involves deciding on a sample rate (e.g. 16kHz), which is analogous to dots-per-inch (dpi) resolution in text scanning and determines the amount of detail captured in digital form. Generally the sample rates possible today are high enough for quality not to be an issue.

  The analog input to the conversion process is a key factor in the quality of the result. In general using the original makes sense, since each analog copy adds distortion, but there can be times when physical deterioration (for example, of a vinyl record) makes using a (tape) copy preferable. The right conversion boards, the right racks, even the right cables all contribute to making as perfect a digital copy as possible. It seems often more like art than science.
- *Segmentation/structuring.* Typically this issue receives little or no attention, but is analogous to the kind of structural definition that makes text-based works useful. For very short sound files (less than five minutes), providing structure may not matter, since a listener might be willing to go through the whole file. For very long (multi-hour) files, structure matters immensely, since no one is likely to listen to it all. The situation for sound files is worse than for a text-equivalent, say a book with no table of contents or index, because there is no easy way to scan a page-size unit, or even to flip through quickly (if the file is big enough).

  Segmentation can be done technically in several different ways. One is to break the recording up into separate files, which can be concentrated back together dynamically. Another is to set breakpoints in the sound file to indicate start points. A third way is to do it via distance (time) into the file. Each of these have advantages and disadvantages that trade off size, intervention, and speed.

  The most difficult decisions are about content: where to create segments, what represents a logical unit for a particular work. Some general rules can be established, such as leaving questions and answers together as a unit, and some speech

Library Hi Tech
Volume 17 · Number 3 · 1999 · 230–232

segments easily because the speakers naturally think and speak in paragraphs (or are actually reading a written speech). But many segmentation decisions need to be done on a case-by-case basis. The cost in time can be high, which is one reason why many digitizers default to creating very large files or just clipping out small bits.

- *Metadata.* These issues can be divided between the bibliographic metadata which apply to the whole work, and structural metadata that apply to individual segments. The former can be handled with standard cataloging, though it is important to include some specialized information, such as sampling rate, format, length (in time), and file size. These metadata can be dissociated from the digital object (as a separate MARC record in an OPAC), or associated with it via virtual HTML (using Dublin Core) or SGML (using TEI) envelopes. Much of it can also go in .wav headers that are part of the sound file itself.

  The structural metadata are necessary to describe particular segments of a work. It can be as short as a normal table of contents entry or as complete as a bibliographic citation. A lot depends on the nature of the segment. Those that resemble separate articles in a journal (a series of speeches at a conference, for example) will need more thorough descriptions than, say, the parts of a two-hour Teddy Roosevelt speech. Either SGML or HTML could be used.
- *Maintenance/refreshing.* The heart of the concern over whether digitization is preservation lies in whether a reliable system can be established for moving digital information from one physical medium to another and for updating the formats as systems change. Fortunately digital audio formats are less version dependent than integrated multimedia products, but the need to be aware of program and operating system dependencies is no less critical. The more standardization there is, the more libraries will be able to share resources for when and how to accomplish this.

  No one should pretend that the costs will be trivial.

A number of important issues do not appear on this list. Among them are integrity (i.e. how to guarantee that a sound file has not been changed) and watermarking (how to brand ownership information onto the sound file). I left them out because they are not preservation issues so much as use issues for digital sound. They will need to be dealt with, too, though perhaps not as part of the same process.

At the past several American Library Association meetings, an informal group has gathered to discuss the best practices for preserving sound in digital formats. This discussion is open to all with interest or expertise. Contact me for more information.

Reference

Pirandello, L. (1985), *Six Characters in Search of an Author*, in Linstrum, J. (Trans.) *Luigi Pirandello: Three Plays*, Methuen, London, p. 85.